

STEROIDAL SAPOGENINS

I. EXTRACTION, ISOLATION, AND IDENTIFICATION

By MONROE E. WALL, MERLE M. KRIDER, EDWARD S. ROTHMAN, AND C. ROLAND EDDY

(From the Eastern Regional Research Laboratory,* Philadelphia, Pennsylvania)

In recent years the steroidal sapogenins have received considerable attention as precursors for the synthesis of sex hormones and cortisone (1-3). Largely through the brilliant researches of Marker and his associates (1), the structure of these compounds is for the most part well established. Unfortunately, data on methods for the isolation and identification of these compounds from natural sources are scanty. The procedure presented by Marker *et al.* lacks detail, is difficult to duplicate, and makes little use of modern techniques (1). Therefore, we are reporting methods perfected during a study of more than 1000 sapogenaceous plant samples.

Isolation of Crude Sapogenin

Sapogenins are found in plant tissues in a combined, glycosidal form called saponins. The method developed involves extraction of saponins with hot 85 to 95 per cent ethanol or isopropanol, removal of fat-soluble material from the filtered, concentrated aqueous alcoholic solution with benzene, extraction of the saponins from the aqueous phase with butanol, and acid hydrolysis to form sapogenins, followed by treatment with hot methanol-potassium hydroxide to remove phenolic and other acidic substances.

We have obtained more consistent results with fresh plant material than with dehydrated tissue, and hence recommend fresh material for the evaluation of the sapogenin content of unknown samples.

The procedure used is illustrated by a typical experiment with a 10 kilo batch of fresh *Agave* leaves.

The fibrous sample is coarsely chopped with a meat cleaver or machete. It is ground several times through a Ball and Jewell¹ grinder provided with

* One of the laboratories of the Bureau of Agricultural and Industrial Chemistry, Agricultural Research Administration, United States Department of Agriculture. This work was done as part of a cooperative arrangement between the Bureau of Plant Industry, Soils, and Agricultural Engineering and Bureau of Agricultural and Industrial Chemistry, United States Department of Agriculture, and the National Institutes of Health, Federal Security Administration.

¹ Mention of a trade name does not imply recommendation or endorsement by the United States Department of Agriculture of this brand over others not mentioned.

a 1 inch screen. The ground material is placed at once in a 50 gallon, steam-jacketed kettle equipped with an air stirrer and having an exit valve at the bottom. Ethanol or isopropanol (85 to 95 per cent) is run in with stirring until the ground plant material is covered by approximately an inch of solvent, forming a mixture which can be easily stirred. The alcohol is then refluxed for 45 minutes with constant stirring. The suspension is cooled and filtered through a canvas filter. A second charge of solvent is added and the process repeated. After the second extraction, the residual meal is allowed to dry in air overnight and is then reground to pass through a 1/16 inch screen. The finely ground material is then extracted a third time. All the alcohol extracts are combined and concentrated at atmospheric pressure to 2 to 4 liters in a 10 gallon steam-jacketed still.

The concentrated extract obtained from extraction of wet plant material usually contains less than 50 per cent alcohol by volume. On standing, considerable precipitation of colloidal material takes place. The extracts are therefore heated and filtered on a Büchner funnel, with large quantities of filter aid. The resultant filter cake containing insoluble material is then stirred with boiling 50 per cent ethanol and filtered. The filter cake is washed several times, and the clarified washes and filtrates are combined.

The extract is freed of interfering fat-soluble material and plant pigments by extraction with benzene. The "defatting" is best carried out in a continuous liquid-liquid extractor in order to avoid emulsions. Under our experimental conditions, 2 to 3 liters of benzene saturated with aqueous alcohol are placed in a boiling flask and continuously passed through a fritted disk of medium porosity immersed in the aqueous layer. The extraction process required 4 to 8 hours.

After the benzene extraction, the benzene is washed once with 1 liter of 50 per cent ethanol (separatory funnel), and the alcoholic aqueous layer is combined with the main benzene-extracted alcoholic fraction. The benzene is discarded.

The benzene-extracted aqueous alcohol solution is concentrated to remove as much alcohol as possible and yet remain free flowing. To the rather turbid aqueous solution are added 5 gm. of sodium chloride per 100 ml. of extract and sufficient HCl to make the pH 4.0 to 5.0. These conditions favor the transfer of saponins into the butanol phase. The extract thus obtained is shaken (separatory funnel) four times with butanol saturated with water; 0.5 volume of butanol to 1 volume of extract is used for each extraction. The butanol layers are combined and washed once with 1 liter of 5 per cent aqueous sodium chloride solution. The washings are extracted with 0.5 liter of butanol. The aqueous layers are then discarded. A quantity of water equal to half the volume of butanol is added to the butanol-saponin extract. The two-phase system is concentrated at atmos-

pheric pressure to a volume of 500 to 1500 ml. Under these conditions, the butanol is driven off in a constant boiling mixture and can be reused. To the aqueous, turbid, saponin solution is added sufficient 95 per cent ethanol so that the final solution is 25 per cent ethanol by volume. Addition of alcohol increases the solubility of the saponins and also reduces foaming during the subsequent hydrolysis step.

Sufficient technical, concentrated hydrochloric acid is added to the aqueous saponin solution to make it 4 N. The solution is refluxed for 3 to 4 hours, cooled, and filtered, and the crude, tarry sapogenin precipitate is washed with 50 per cent aqueous ethanol.

The precipitate is refluxed for 1 hour with a mixture of 3 liters of benzene and 1 liter of methanol containing 200 gm. of potassium hydroxide. The solution is cooled and filtered, and any residue is washed with hot benzene

TABLE I
Butanol Extraction of Steroidal Saponins from Aqueous Solutions
System, 5 ml. of butanol and 5 ml. of water, mutually saturated.

Saponin	Quantity	Per cent in butanol				Total recovery
		Extraction 1	Extraction 2	Extraction 3	Extraction 4	
	mg.					per cent
Digitonin	109	97	1.4	0.2	0.1	98.7
Dioscin	113	94	2.2	0.9	0.3	97.4
Gitonin	108	86.5	8.2	2.5	2.0	99.2
Sarsasaponin	83	70.5	17.8	4.7	1.0	94.0

containing 10 per cent ethanol. The filtrate and washings are combined, 1 liter of water is added, and the aqueous layer is drawn off. The aqueous layer is then twice reextracted with 1 liter of benzene each time. All the benzene solutions are combined and concentrated to dryness. The residual, usually tan, solids are a mixture of one or more sapogenins with about 50 per cent resinous material.

Since most sapogenins are quite insoluble in cold acetone, methanol, or hexane, much resinous material can be removed by triturating the finely ground crude sapogenin with one of these solvents and filtering. The procedure and solvent vary from sample to sample and must be used with caution. The filtrate should always be evaporated and examined for the presence of appreciable quantities of sapogenins by means of crystallization and chromatography.

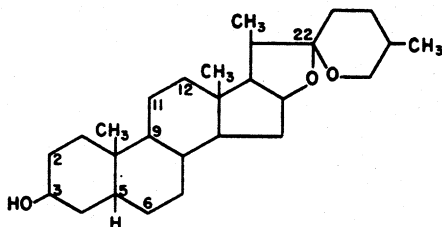
The procedure described above has been carefully checked. It has been found that the first alcoholic extraction removes 75 to 85 per cent of the total saponins extracted (based on sapogenin assay). The second and

third extractions account for the residual material. Further extractions yield negligible quantities. Similar results were found with *Dioscorea*, *Yucca*, and *Agave* species, the most common sapogenaceous plants. Small scale analytical extraction studies gave the same type of data as large scale experiments.

The extraction of saponins from aqueous solution with butanol has proved quite satisfactory. The solution must be neutral or acid, since extraction from alkaline media gives poor yields. The butanol extraction separates saponins from sugars and proteins and thus markedly reduces tar and resin formation in the subsequent acid hydrolysis. Table I shows that water-saturated butanol will quantitatively remove pure saponins from aqueous solution.

Separation of Sapogenins

Structural Differences—More than 25 steroidal sapogenins have been described (1, 2, 4). The structure of smilagenin, a typical example, is shown



SMILAGENIN

in the diagram. The most common variants are ring inversion at C₂₂, cis or trans hydrogen isomerism at C₅, unsaturation between C₅ and C₆ or between C₉ and C₁₁, presence of carbonyl at C₁₂, and additional hydroxyl groups at C₂, C₆, or C₁₂.

Previous separations of sapogenins have been based principally on differences in solubility of the various sapogenins or their acetates in solvents such as ether, acetone, methanol, or acetic anhydride (1). These procedures were carried out with large quantities (100 to 500 gm.) of sapogenins. We have been unable to duplicate these results with quantities of total sapogenins ranging from 1 to 25 gm.

Chromatographic Procedure—Merck or Alcoa activated alumina was used as an adsorbent. 10 to 20 times as much adsorbent as sapogenin were used to eliminate possible break through.

The crude sapogenin is stirred with 500 to 1000 ml. of hot benzene containing 2 per cent chloroform. The solution is cooled and filtered, and the filtrate is then placed on the column. Any residual precipitate is washed with more benzene-2 per cent chloroform and is passed through the column.

Elution with benzene-2 per cent chloroform is then continued in the usual manner until no more sapogenin is eluted (Fraction 1). The chloroform concentration is then increased to 20 per cent (Fraction 2), and finally 20 per cent ethanol in benzene (Fraction 3) is used. The eluent is collected in 250 ml. fractions which are evaporated to dryness. Melting points are taken, and appropriate preliminary combinations of fractions are made.

In this manner three fractions consisting respectively of monohydroxy non-ketonic (Fraction 1), monohydroxy ketonic (Fraction 2), and dihydroxy sapogenins (Fraction 3) are separated. To separate dihydroxy ketonic from non-ketonic sapogenins, it is necessary to convert Fraction 3 to the acetate. Alumina or preferably Florisil is used as the adsorbent. The acetates are taken up in hexane containing 20 per cent benzene. Non-ketonic sapogenin acetates are eluted with this mixture and ketones with benzene or benzene-chloroform mixtures.

Identification of Sapogenins

Prior to further identification, the various chromatographic fractions are combined on the basis of rough melting points and adsorption behavior. Monohydroxy non-carbonyl compounds usually melt below 210°. These are crystallized from methanol or from methanol-benzene if difficultly soluble in boiling methanol. Almost all sapogenins that melt above 210° are ketonic or dihydroxy. These are best crystallized from ether-methanol.

Acetates are made routinely on all fractions by refluxing the genins for 1 hour with an excess of acetic anhydride containing a few drops of pyridine. The solution is cooled, and the genin acetate is filtered off and recrystallized from methanol.

Melting Point—The melting points of the genins and their acetates were determined with a Kofler microscopic melting point apparatus having polarizing disks. Not only can the melting points be determined accurately, but also the crystal form or habit can be observed. At times this latter property is useful in identifying the compound. Mixed melting points with known compounds can be used as a tentative aid in identification (1). The method is none too reliable, particularly if the unknown is even slightly impure.

Specific Rotation—The specific rotations in chloroform of all genins and their acetates were determined. The rotations were determined at 25° (sodium lamp) at a concentration of 6 to 10 mg. per ml. 2 and 4 dm. micro polarimeter tubes were used.

The specific rotations of the sapogenins and their acetates are properties useful in identification. In some cases the two rotations constitute a unique means of identification. In other cases, particularly if the sapogenin has been given a preliminary group classification (monohydroxy,

dihydroxy, etc.), rotations can be used to determine specific features of the molecule.

For example, unsaturation in sapogenins can be determined more easily by specific rotation than by any other method, the infra-red absorption of an ethylenic bond being extremely weak (5). Table II shows that the rotation of diosgenin is much more negative than that of any of its monohydroxy saturated analogues. Similarly, compare yuccagenin with gito-genin, kammogenin with manogenin, 9-dehydrohecogenin with hecogenin.

Infra-Red Data—Infra-red absorption spectra of the sapogenins and their acetates are of great value in characterizing sapogenins. A Beckman model IR-3 infra-red spectrophotometer¹ was used in our work. The chief uses of the method are as follows:

"Finger-Print" Region—The portion of the spectrum between 700 and 1400 cm^{-1} (14 to 7 μ) can be used to characterize the individual sapogenins, and unknown sapogenins can be identified by comparison with a file of known sapogenin spectra. In this region sapogenins are best characterized as their acetates in carbon disulfide. Concentrations of 10 gm. per liter are used in 1.0 mm. cells. Free sapogenins can also be used in chloroform, at 25 gm. per liter in 0.5 mm. cells, but are less satisfactory. A paper in preparation will give finger-print spectra of the most common sapogenins.

Determination of Isomerism at C_6 —Jones *et al.* (9) have shown that steroid acetates give either single or multiple absorption bands in the region 1200 to 1300 cm^{-1} (8.3 to 7.7 μ), depending on the relative configuration of the 3-hydroxy and 5-hydrogen groups. We find exactly the same effects with sapogenin monoacetates as they found for simpler steroids. A single band at 1242 to 1244 cm^{-1} (8.05 to 8.04 μ) is found for all the trans or unsaturated sapogenin monoacetates listed in Table II and a multiple band for the cis compounds. Since both cis and trans dihydroxysapogenin acetates give a multiple band in this region, the dihydroxy compounds cannot be differentiated with certainty by this method.

Determination of Isomerism at C_{22} —In studies at this laboratory, detailed in a paper in preparation, it has been found that sapogenins with "normal" or "iso" structures at C_{22} can be differentiated. All isosapogenin acetates show strong absorption bands near 865, 900, 920, and 981 cm^{-1} (11.55, 11.1, 10.85, and 10.19 μ), with the 900 band much stronger than the 920 band. The acetate of sarsasapogenin, the only commonly occurring normal sapogenin, has strong bands at 851, 896, 921, and 986 cm^{-1} (11.75, 11.2, 10.86, and 10.14 μ), with the 921 band much stronger than the 896 band. Similar compounds with the terminal sixth ring opened have only weak absorption in this region, indicating that these four bands may be due to the terminal ring only. From the relative strengths of the bands near 900 and 920 cm^{-1} , normal and isosapogenins can be readily distinguished, with confirmation from the positions of the remaining two bands.

TABLE II
Physical Constants of Sapogenins

Compound*	M.p. (Kofler)		Specific rotation,† 25°, chloroform, sodium arc		Configuration and structure		
	Genin	Acetate	Genin	Acetate	C ₁	C ₂	
Monohydroxy, non-ketonic							
Diosgenin	208	202	−121	−123		Iso	Δ ⁴ -3β-Hydroxy
Sarsasapogenin	200	145	−78	−70	Cis	Normal	3β-Hydroxy
Smilagenin	184	150	−73	−60	"	Iso	"
Tigogenin	206	204	−67	−73	Trans	"	"
Monohydroxy ketonic							
Hecogenin	268	245	+10	−1	Trans	Iso	12-Keto-3β-hydroxy
9-Dehydroheco- genin	235†	227†	−11†	−7†	"	"	Δ ⁹ -12-Keto-3β-hydroxy
Dihydroxy, non-ketonic							
Gitogenin	272	243	−75	−98	Trans	Iso	2,3β-Dihydroxy
Chlorogenin	276	155	−45	−38	"	"	3β,6α-Dihydroxy
Yuccagenin	243	178	−120	−139		"	Δ ⁴ -2,3β-Dihydroxy
Samogenin	202	199	−74	−84	Cis	"	2,3β-Dihydroxy
Dihydroxy ketonic							
Manogenin	242	259	−5	−45	Trans	Iso	12-Keto-2,3β-dihydroxy
9-Dehydromano- genin	240†	263†	−16†	−62†	"	"	Δ ⁹ -12-Keto-2,3β-dihydroxy
Kammogenin	245	259	−53	−83		"	Δ ⁴ -12-Keto-2,3β-dihydroxy

* The compounds listed in this table are named according to the nomenclature of Marker *et al.* (1). Several new nomenclature systems have recently appeared (6, 7). Since the question of naming sapogenins is still not completely resolved, we prefer to use the widely known older system.

† Most of the rotations are from our experimental data, the solvent being chloroform. The rotations followed by † were taken from the data of Wagner *et al.* (8). These were obtained with dioxane as the solvent. Dioxane gives a specific rotation approximately 10° more positive than chloroform.

Determination of Carbonyl-Containing Sapogenins—Unconjugated carbonyl and α,β -unsaturated carbonyl are readily determined in sapogenins in the 1700 cm^{-1} (6μ) region of the spectrum. Jones *et al.* (10–12) found that the frequency of the carbonyl-stretching vibration is a function of its position in the steroid molecule and of the nature of the near-by functional groups. Of the possible carbonyl types, only two are of importance in plant sapogenins: an unconjugated ketone group on C_{12} and a 12-ketone group conjugated to a $\Delta^9, 11$ olefinic group. In agreement with the findings of Jones *et al.* for simpler steroids, we find that ketosapogenins have a single absorption band in the range 1705 to 1715 cm^{-1} (5.87 to 5.83μ) and that α,β -unsaturated ketosapogenins have two bands at 1600 to 1605 cm^{-1} (6.25 to 6.23μ) and 1673 to 1679 cm^{-1} (5.98 to 5.96μ).

The carbonyl determination is carried out on the sapogenins in chloroform at 15 to 25 gm. per liter in a 0.5 mm. cell . Acetates of the sapogenins should not be used because of interference from the acetate carbonyl band. It is essential that chloroform from the same bottle be used as a blank, because different batches of A. C. S. grade chloroform sometimes have different transmissions in this region of the spectrum. We have used hecogenin and 9-dehydrohecogenin as reference substances. Provisional values of their molar absorptivities² are 424 for hecogenin at 1705 cm^{-1} and 468 for 9-dehydrohecogenin at 1675 cm^{-1} . Unknowns can be analyzed by comparing their molar absorptivities with those of these reference substances. The 1600 to 1605 cm^{-1} band is too weak and variable to be suitable for quantitative analysis of unknowns. For mixtures containing carbonyl bands at more than one frequency more accurate analyses can be obtained by using integrated intensities rather than maximal absorptivities (13).

Conversion to Known Compounds—This is rarely a necessity. As a precautionary check, 100 to 200 mg. samples of all carbonyl unknowns are subjected to Wolff-Kishner reduction by the method of Huang-Minlon (14). Under the experimental conditions, hecogenin and manogenin are reduced in excellent yields to tigogenin and gitogenin. Under these conditions, 11-keto compounds should not react or should give poor yields (14). Hydrogenation, oxidation with chromic acid (1), and reduction with sodium (1) and lithium aluminum hydride (15) are also occasionally useful for confirmatory purposes.

Other Techniques—Ultraviolet absorption is useful for determining α,β -unsaturated ketones because of the characteristic strong band (molar absorptivity, $\epsilon = 11,000$) at $238\text{ m}\mu$. A non-conjugated carbonyl group has only a weak maximum (ϵ about 50) at $285\text{ m}\mu$, and this has proved so sub-

² Molar absorptivity (or molar extinction coefficient) = absorbance (or optical density) divided by cell thickness in cm. and by concentration in moles per liter.

ject to interference that the ultraviolet method was discarded as a means of carbonyl assay.

Carbon and hydrogen determinations are rarely required unless the presence of a new or unusual sapogenin is suspected. A few spot checks with carefully purified sapogenins have given results that agreed well with theoretical values.

Correlation of Data

Table II gives melting points, specific rotations, and configuration data of the sapogenins studied. A number of other sapogenins have been described (1, 2, 4), most of which are rare.

Since, in most cases, the various properties of the sapogenins and their acetates can be correlated, at least a tentative identification can be made rapidly. It should be emphasized that relatively small quantities of other sapogenins or non-sapogenin substances will markedly affect melting points, but usually will not greatly alter other physical properties. Slight impurities can lower melting points from 5° to 15°.

A brief description of the data used to identify rapidly the most common sapogenins is presented below.

As described previously, on the basis of adsorption behavior, the genins can be divided into monohydroxy non-ketonic, monohydroxy ketonic, dihydroxy non-ketonic and dihydroxy ketonic. Trihydroxy compounds are rare but would behave like the dihydroxy compounds.

Monohydroxy Non-Ketonic—Absence of the carbonyl group can be confirmed by infra-red spectrophotometry. Diosgenin and tigogenin are sharply differentiated from sarsasapogenin and smilagenin by the melting points of their acetates. Diosgenin is well differentiated from tigogenin by specific rotation. Sarsasapogenin and smilagenin are best differentiated by infra-red determination of C₂₂ configuration.

Monohydroxy Ketonic—The presence of carbonyl can be confirmed by infra-red absorption. Hecogenin and its 9-dehydro analogue are the only examples we have found in this classification. The 9-dehydro component is usually found in small quantity along with the main component hecogenin. It is easily detected by infra-red or ultraviolet spectrophotometry. We have never found botogenin (16) in our *Dioscorea* samples. If present, the compound could be differentiated from hecogenin by its specific rotation, which would undoubtedly be more negative because of its Δ^5 unsaturation. A slightly positive specific rotation is almost a sure indication of hecogenin.

Dihydroxy Non-Ketonic—Absence of carbonyl can be shown by infra-red spectrophotometry. Yuccagenin is identified by the unusually low melting point of its acetate and high negative specific rotation. Gitogenin and

chlorogenin are sharply differentiated by the differences in the melting points of their acetates. Samogenin is distinguished from other dihydroxy genins by the low melting point of its genin. The configuration at C₂₂, determined by infra-red absorption, would help differentiate it from some of its isomers (1) which we have not yet encountered.

Dihydroxy Ketonic—The presence of carbonyl is confirmed by infra-red measurements. Manogenin is sharply differentiated from kammogenin by rotation differences. 9-Dehydromanogenin is usually a trace contaminant in manogenin. It is identified in the same manner as 9-dehydrohecogenin.

SUMMARY

Methods for extraction, isolation, and identification of steroidal saponins have been presented. The saponin precursors are extracted from wet plant tissue with 90 to 95 per cent ethanol. The concentrated extract is freed of fats and lipides by benzene extraction. Saponins are extracted with butanol from the aqueous residue, and, after concentration, hydrolyzed to sapogenins with 4 N hydrochloric acid. The sapogenins are extracted from the hydrolysis mixture with benzene and further purified by hot methanol-potassium hydroxide and ether trituration. The purified sapogenins are adsorbed on alumina and resolved into four groups: monohydroxy non-ketonic, monohydroxy ketonic, dihydroxy non-ketonic, and dihydroxy ketonic. The sapogenins thus isolated are further identified by melting points, specific rotations, and infra-red spectra.

We appreciate the encouragement and advice offered by J. J. Willaman, B. A. Brice, and G. C. Nutting during this investigation. We are indebted to A. Turner for ultraviolet observations, L. P. Witnauer for x-ray data, and C. O. Willits and C. L. Ogg for carbon and hydrogen analyses. We thank J. W. Garvin, H. I. Sinnamon, J. R. Willis, M. E. Klumpp, M. A. Morris, and J. R. Necho for assistance in various phases of this investigation. We wish to thank T. D. Fontaine for supplying us with samples of dihydrosapogenins.

BIBLIOGRAPHY

1. Marker, R. E., Wagner, R. B., Ulshafer, P. R., Wittbecker, E. L., Goldsmith, D. P. J., and Ruof, C. H., *J. Am. Chem. Soc.*, **69**, 2167 (1947).
2. Marker, R. E., and Applezweig, N., *Chem. Eng. News*, **27**, 3348 (1949).
3. Rosenkranz, G., Pataki, J., and Djerassi, C., *J. Am. Chem. Soc.*, **73**, 4055 (1951).
4. Fieser, L. F., and Fieser, M., Natural products related to phenanthrene, American Chemical Society monograph series, New York, 3rd edition (1949).
5. Jones, R. N., Humphries, P., Packard, E., and Dobriner, K., *J. Am. Chem. Soc.*, **72**, 86 (1950).
6. Rosenkranz, G., and Djerassi, C., *Nature*, **166**, 104 (1950).
7. Vorschläge zur Nomenklatur der Steroide, *Helv. chim. acta*, **34**, 1680 (1951).

8. Wagner, R. B., Forker, R. F., and Spitzer, P. F., Jr., *J. Am. Chem. Soc.*, **73**, 2494 (1951).
9. Jones, R. N., Humphries, P., Herling, F., and Dobriner, K., *J. Am. Chem. Soc.*, **73**, 3215 (1951).
10. Jones, R. N., Humphries, P., and Dobriner, K., *J. Am. Chem. Soc.*, **71**, 241 (1949).
11. Jones, R. N., Humphries, P., and Dobriner, K., *J. Am. Chem. Soc.*, **72**, 956 (1950).
12. Jones, R. N., Williams, V. Z., Whalen, M. J., and Dobriner, K., *J. Am. Chem. Soc.*, **70**, 2024 (1948).
13. Jones, R. N., Ramsay, D. A., Keir, D. S., and Dobriner, K., *J. Am. Chem. Soc.*, **74**, 80 (1952).
14. Huang-Minlon, *J. Am. Chem. Soc.*, **71**, 3301 (1949).
15. Djerassi, C., Martinez, H., and Rosenkranz, G., *J. Org. Chem.*, **16**, 1278 (1951).
16. Marker, R. E., *J. Am. Chem. Soc.*, **71**, 2656 (1949).